

# Infimal Convolution Regularisation Functionals of BV and $L^p$ Spaces. The Case $p = \infty$

Martin Burger<sup>1</sup>, Konstantinos Papafitsoros<sup>2</sup>, Evangelos Papoutsellis<sup>3</sup>, and Carola-Bibiane Schönlieb<sup>3</sup>

<sup>1</sup> Institute for Computational and Applied Mathematics, University of Münster, Germany

<sup>2</sup> Institute for Mathematics, Humboldt University of Berlin, Germany

<sup>3</sup> Department of Applied Mathematics and Theoretical Physics, University of Cambridge, United Kingdom

`martin.burger@wwu.de`, `papafitsoros@hu-berlin.de`, `{ep374,cbs31}@cam.ac.uk`

**Abstract.** In this paper we analyse an infimal convolution type regularisation functional called  $\text{TVL}^\infty$ , based on the total variation (TV) and the  $L^\infty$  norm of the gradient. The functional belongs to a more general family of  $\text{TVL}^p$  functionals ( $1 < p \leq \infty$ ) introduced in [5]. There, the case  $1 < p < \infty$  is examined while here we focus on the  $p = \infty$  case. We show via analytical and numerical results that the minimisation of the  $\text{TVL}^\infty$  functional promotes piecewise affine structures in the reconstructed images similar to the state of the art total generalised variation (TGV) but improving upon preservation of hat-like structures. We also propose a spatially adapted version of our model that produces results comparable to TGV and allows space for further improvement.

**Keywords:** Total Variation, Infimal Convolution,  $L^\infty$  norm, Denoising, Staircasing

## 1 Introduction

In the variational setting for imaging, given image data  $f \in L^s(\Omega)$ ,  $\Omega \subseteq \mathbb{R}^2$ , one aims to reconstruct an image  $u$  by minimising a functional of the type

$$\min_{u \in X} \frac{1}{s} \|f - Tu\|_{L^s(\Omega)}^s + \Psi(u), \quad (1.1)$$

over a suitable function space  $X$ . Here  $T$  denotes a linear and bounded operator that encodes the transformation or degradation that the original image has gone through. Random noise is also usually present in the degraded image, the statistics of which determine the norm in the first term of (1.1), the *fidelity term*. The presence of  $\Psi$ , the *regulariser*, makes the minimisation (1.1) a well-posed problem and its choice is crucial for the overall quality of the reconstruction. A classical regulariser in imaging is the total variation functional weighted with a parameter  $\alpha > 0$ ,  $\alpha \text{TV}$  [10], where

$$\text{TV}(u) := \sup \left\{ \int_{\Omega} u \operatorname{div} \phi \, dx : \phi \in C_c^1(\Omega, \mathbb{R}^2), \|\phi\|_{\infty} \leq 1 \right\}. \quad (1.2)$$

While TV is able to preserve edges in the reconstructed image, it also promotes piecewise constant structures leading to undesirable staircasing artefacts. Several regularisers that incorporate higher order derivatives have been introduced in order to resolve this issue. The most prominent one has been the second order total generalised variation (TGV) [3] which can be interpreted as a special type of infimal convolution of first and second order derivatives. Its definition reads

$$\text{TGV}_{\alpha,\beta}^2(u) := \min_{w \in \text{BD}(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|\mathcal{E}w\|_{\mathcal{M}}. \quad (1.3)$$

Here  $\alpha, \beta$  are positive parameters,  $\text{BD}(\Omega)$  is the space of functions of bounded deformation,  $\mathcal{E}w$  is the distributional symmetrised gradient and  $\|\cdot\|_{\mathcal{M}}$  denotes the Radon norm of a finite Radon measure, i.e.,

$$\|\mu\|_{\mathcal{M}} = \sup \left\{ \int_{\Omega} \phi d\mu : \phi \in C_c^\infty(\Omega, \mathbb{R}^2), \|\phi\|_\infty \leq 1 \right\}.$$

TGV regularisation typically produces piecewise smooth reconstructions eliminating the staircasing effect. A plausible question is whether results of similar quality can be achieved using simpler, first order regularisers. For instance, it is known that Huber TV can reduce the staircasing up to an extent [6].

In [5], a family of first order infimal convolution type regularisation functionals is introduced, that reads

$$\text{TVL}_{\alpha,\beta}^p(u) := \min_{w \in L^p(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^p(\Omega)}, \quad (1.4)$$

where  $1 < p \leq \infty$ . While in [5], basic properties of (1.4) are shown for the general case  $1 < p \leq \infty$ , see Proposition 1, the main focus remains on the finite  $p$  case. There, the  $\text{TVL}^p$  regulariser is successfully applied to image denoising and decomposition, reducing significantly the staircasing effect and producing piecewise smooth results that are very similar to the solutions obtained by TGV. Exact solutions of the  $L^2$  fidelity denoising problem are also computed there for simple one dimensional data.

### Contribution of the Present Work

The purpose of the present paper is to examine more thoroughly the case  $p = \infty$ , i.e.,

$$\text{TVL}_{\alpha,\beta}^\infty(u) := \min_{w \in L^\infty(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^\infty(\Omega)}, \quad (1.5)$$

and the use of the  $\text{TVL}^\infty$  functional in  $L^2$  fidelity denoising

$$\min_u \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \text{TVL}_{\alpha,\beta}^\infty(u). \quad (1.6)$$

We study thoroughly the one dimensional version of (1.6), by computing exact solutions for data  $f$  a piecewise constant and a piecewise affine step function. We show that the solutions are piecewise affine and we depict some similarities and differences to TGV solutions. The functional  $\text{TVL}_{\alpha,\beta}^\infty$  is further tested for Gaussian denoising. We show that  $\text{TVL}_{\alpha,\beta}^\infty$ , unlike TGV, is able to recover hat-like structures, a property that is already present in the  $\text{TVL}_{\alpha,\beta}^p$  regulariser for large values of  $p$ , see [5], and it is enhanced here. After explaining some

limitations of our model, we propose an extension where the parameter  $\beta$  is spatially varying, i.e.,  $\beta = \beta(x)$ , and discuss a rule for selecting its values. The resulting denoised images are comparable to the TGV reconstructions and indeed the model has the potential to produce much better results.

## 2 Properties of the $\text{TVL}_{\alpha,\beta}^\infty$ Functional

The following properties of the  $\text{TVL}_{\alpha,\beta}^\infty$  functional are shown in [5]. We refer the reader to [5,9] for the corresponding proofs and to [1] for an introduction to the space of functions of bounded variation  $\text{BV}(\Omega)$ .

**Proposition 1 ([5])** *Let  $\alpha, \beta > 0$ ,  $d \geq 1$ , let  $\Omega \subseteq \mathbb{R}^d$  be an open, connected domain with Lipschitz boundary and define for  $u \in L^1(\Omega)$*

$$\text{TVL}_{\alpha,\beta}^\infty(u) := \min_{w \in L^\infty(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^\infty(\Omega)}. \quad (2.1)$$

*Then we have the following:*

- (i)  $\text{TVL}_{\alpha,\beta}^\infty(u) < \infty$  if and only if  $u \in \text{BV}(\Omega)$ .
- (ii) If  $u \in \text{BV}(\Omega)$  then the minimum in (2.1) is attained.
- (iii)  $\text{TVL}_{\alpha,\beta}^\infty(u)$  can equivalently be defined as

$$\text{TVL}_{\alpha,\beta}^\infty(u) = \sup \left\{ \int_{\Omega} u \operatorname{div} \phi \, dx : \phi \in C_c^1(\Omega, \mathbb{R}^2), \|\phi\|_\infty \leq \alpha, \|\phi\|_{L^1(\Omega)} \leq \beta \right\},$$

*and  $\text{TVL}_{\alpha,\beta}^\infty$  is lower semicontinuous w.r.t. the strong  $L^1$  convergence.*

- (iv) *There exist constants  $0 < C_1 < C_2 < \infty$  such that*

$$C_2 \text{TV}(u) \leq \text{TVL}_{\alpha,\beta}^\infty(u) \leq C_1 \text{TV}(u), \quad \text{for all } u \in \text{BV}(\Omega).$$

- (v) *If  $f \in L^2(\Omega)$ , then the minimisation problem*

$$\min_{u \in \text{BV}(\Omega)} \frac{1}{s} \|f - u\|_{L^2(\Omega)}^2 + \text{TVL}_{\alpha,\beta}^\infty(u),$$

*has a unique solution.*

## 3 The One Dimensional $L^2$ - $\text{TVL}_{\alpha,\beta}^\infty$ Denoising Problem

In order to get an intuition about the underlying regularising mechanism of the  $\text{TVL}_{\alpha,\beta}^\infty$  regulariser, we study here the one dimensional  $L^2$  denoising problem

$$\min_{\substack{u \in \text{BV}(\Omega) \\ w \in L^\infty(\Omega)}} \frac{1}{2} \|f - u\|_{L^2(\Omega)}^2 + \alpha \|Du - w\|_{\mathcal{M}} + \beta \|w\|_{L^\infty(\Omega)}. \quad (3.1)$$

In particular, we present exact solutions for simple data functions. In order to do so, we use the following theorem:

**Theorem 2 (Optimality conditions)** *Let  $f \in L^2(\Omega)$ . A pair  $(u, w) \in \text{BV}(\Omega) \times L^\infty(\Omega)$  is a solution of (3.1) if and only if there exists a unique  $\phi \in H_0^1(\Omega)$  such*

that

$$\phi' = u - f, \quad (3.2)$$

$$\phi \in \alpha \text{Sgn}(Du - w), \quad (3.3)$$

$$\phi \in \begin{cases} \{\psi \in L^1(\Omega) : \|\psi\|_{L^1(\Omega)} \leq \beta\}, & \text{if } w = 0, \\ \{\psi \in L^1(\Omega) : \langle \psi, w \rangle = \beta \|w\|_{L^\infty(\Omega)}, \|\psi\|_{L^1(\Omega)} \leq \beta\}, & \text{if } w \neq 0. \end{cases} \quad (3.4)$$

Recall that for a finite Radon measure  $\mu$ ,  $\text{Sgn}(\mu)$  is defined as

$$\text{Sgn}(\mu) = \left\{ \phi \in L^\infty(\Omega) \cap L^\infty(\Omega, \mu) : \|\phi\|_{L^\infty(\Omega)} \leq 1, \phi = \frac{d\mu}{d|\mu|}, |\mu| - \text{a.e.} \right\}.$$

As it is shown in [4],  $\text{Sgn}(\mu) \cap C_0(\Omega) = \partial \|\cdot\|_{\mathcal{M}}(\mu) \cap C_0(\Omega)$ .

*Proof.* The proof of Theorem 2 is based on Fenchel–Rockafellar duality theory and follows closely the corresponding proof of the finite  $p$  case. We thus omit it and we refer the reader to [5,9] for further details, see also [4,8] for the analogue optimality conditions for the one dimensional  $L^1$ –TGV and  $L^2$ –TGV problems.  $\square$

The following proposition states that the solution  $u$  of (3.1) is essentially piecewise affine.

**Proposition 3 (Affine structures)** *Let  $(u, w)$  be an optimal solution pair for (3.1) and  $\phi$  be the corresponding dual function. Then  $|w| = \|w\|_{L^\infty(\Omega)}$  a.e. in the set  $\{\phi \neq 0\}$ . Moreover,  $|u'| = \|w\|_{L^\infty(\Omega)}$  whenever  $u > f$  or  $u < f$ .*

*Proof.* Suppose that there exists a  $U \subseteq \{\phi \neq 0\}$  of positive measure such that  $|w(x)| < \|w\|_{L^\infty(\Omega)}$  for every  $x \in U$ . Then

$$\begin{aligned} \int_{\Omega} \phi w \, dx &\leq \int_{\Omega \setminus U} |\phi| |w| \, dx + \int_U |\phi| |w| \, dx < \|w\|_{L^\infty(\Omega)} \left( \int_{\Omega \setminus U} |\phi| \, dx + \int_U |\phi| \, dx \right) \\ &= \|w\|_{L^\infty(\Omega)} \|\phi\|_{L^1(\Omega)} = \beta \|w\|_{L^\infty(\Omega)}, \end{aligned}$$

where we used the fact that  $\|\phi\|_{L^1(\Omega)} \leq \beta$  from (3.4). However this contradicts the fact that  $\langle \phi, w \rangle = \beta \|w\|_{L^\infty(\Omega)}$  also from (3.4). Note also that from (3.2) we have that  $\{u > f\} \cup \{u < f\} \subseteq \{\phi \neq 0\}$  up to null sets. Thus, the last statement of the proposition follows from the fact that whenever  $u > f$  or  $u < f$  then  $u' = w$  there. This last fact can be shown exactly as in the corresponding TGV problems, see [4, Prop. 4.2].  $\square$

Piecewise affinity is typically a characteristic of higher order regularisation models, e.g. TGV. Indeed, as the next proposition shows, TGV and  $\text{TVL}^\infty$  regularisation coincide in some simple special cases.

**Proposition 4** *The one dimensional functionals  $\text{TGV}_{\alpha,\beta}^2$  and  $\text{TVL}_{\alpha,2\beta}^\infty$  coincide in the class of those BV functions  $u$ , for which an optimal  $w$  in both definitions of TGV and  $\text{TVL}^\infty$  is odd and monotone.*

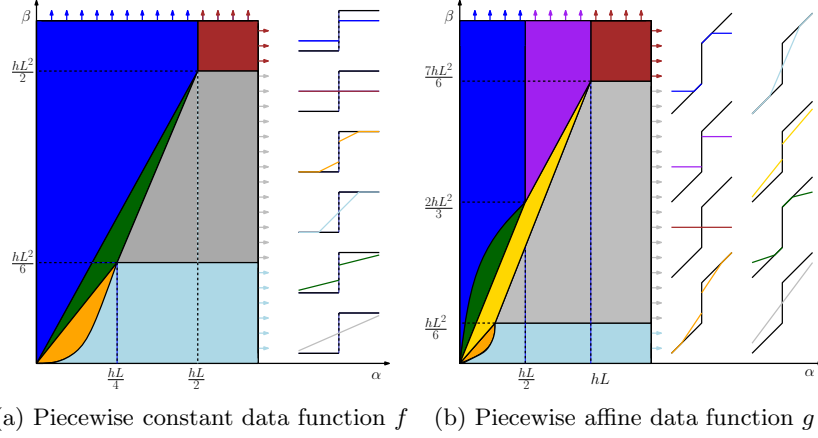


Fig. 1: Exact solutions for the  $L^2$ - $\text{TVL}^\infty_{\alpha,\beta}$  one dimensional denoising problem

*Proof.* Note first that for every odd and monotone bounded function  $w$  we have  $\|Dw\|_{\mathcal{M}} = 2\|w\|_{L^\infty(\Omega)}$  and denote this set of functions by  $\mathcal{A} \subseteq \text{BV}(\Omega)$ . For a BV function  $u$  as in the statement of the proposition we have

$$\begin{aligned} \text{TGV}^2_{\alpha,\beta}(u) &= \argmin_{w \in \mathcal{A}} \alpha \|Du - w\|_{\mathcal{M}} + \beta \|Dw\|_{\mathcal{M}} \\ &= \argmin_{w \in \mathcal{A}} \alpha \|Du - w\|_{\mathcal{M}} + 2\beta \|w\|_{L^\infty(\Omega)} = \text{TVL}^\infty_{\alpha,2\beta}(u). \end{aligned}$$

□

**Exact solutions:** We present exact solutions for the minimisation problem (3.1), for two simple functions  $f, g : (-L, L) \rightarrow \mathbb{R}$  as data, where  $f(x) = h\mathcal{X}_{(0,L)}(x)$  and  $g(x) = f(x) + \lambda x$ , with  $\lambda, h > 0$ . Here  $\mathcal{X}_C(x) = 1$  for  $x \in C$  and 0 otherwise.

With the help of the optimality conditions (3.2)–(3.4) we are able to compute all possible solutions of (3.1) for data  $f$  and  $g$  and for all values of  $\alpha$  and  $\beta$ . These solutions are depicted in Figure 1. Every coloured region corresponds to a different type of solution. Note that there are regions where the solutions coincide with the corresponding solutions of TV minimisation, see the blue and red regions in Figure 1a and the blue, purple and red regions in Figure 1b. This is not surprising since as it is shown in [5] for all dimensions,  $\text{TVL}^\infty_{\alpha,\beta} = \alpha \text{TV}$ , whenever  $\beta/\alpha \geq |\Omega|^{1/q}$  with  $1/p + 1/q = 1$  and  $p \in (1, \infty]$ . Notice also the presence of affine structures in all solutions, predicted by Proposition 3. For demonstration purposes, we present the computation of the exact solution that corresponds to the yellow region of Figure 1b and refer to [9] for the rest.

Since we require a piecewise affine solution, from symmetry and (3.2) we have that  $\phi(x) = (c_1 - \lambda)\frac{x^2}{2} - c_2|x| + c_3$ . Since we require  $u$  to have a discontinuity at 0, (3.3) implies  $\phi(0) = a$  while from the fact that  $\phi \in H^1_0(\Omega)$  and from (3.4) we must have  $\phi(-L) = 0$  and  $\langle \phi, w \rangle = \beta \|w\|_{L^\infty(\Omega)}$ . These conditions give

$$c_1 = \frac{6(\alpha L - \beta)}{L^3} + \lambda, \quad c_2 = \frac{4\alpha L - 3\beta}{L^2}, \quad c_3 = \alpha.$$

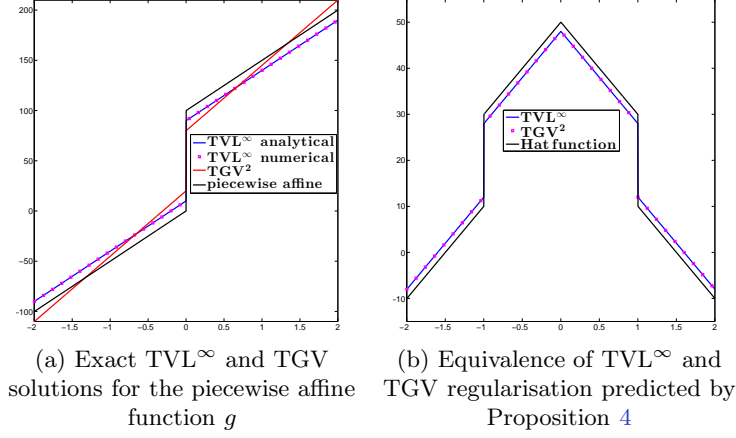


Fig. 2: One dimensional numerical examples

We also have  $c_1 = u' = w$  and thus we require  $c_1 > 0$ . Since we have a jump at  $x = 0$  we also require  $g(0) < u(0) < h$ , i.e.,  $0 < c_2 < \frac{h}{2}$  and  $u(-L) > g(-L)$  i.e.,  $\phi'(-L) > 0$ . These last inequalities are translated to the following conditions

$$\left\{ \beta < \alpha L + \frac{\lambda L^3}{6}, \beta > \frac{4\alpha L}{3} - \frac{hL^2}{6}, \beta > \frac{2\alpha L}{3}, \beta < \frac{4\alpha L}{3} \right\},$$

which define the yellow area in Figure 1b. We can easily compute  $u$  now:

$$u(x) = \begin{cases} \left( \frac{6(\alpha L - \beta)}{L^3} + \lambda \right) x + h - \frac{4\alpha L - 3\beta}{L^2}, & x \in (0, L), \\ \left( \frac{6(\alpha L - \beta)}{L^3} + \lambda \right) x + \frac{4\alpha L - 3\beta}{L^2}, & x \in (-L, 0). \end{cases}$$

Observe that when  $\beta = \alpha L$ , apart from the discontinuity, we can also recover the slope of the data  $g' = \lambda$ , something that neither TV nor TGV regularisation can give for this example, see [8, Section 5.2].

## 4 Numerical Experiments

In this section we present our numerical experiments for the discretised version of  $L^2\text{-TVL}_{\alpha,\beta}^\infty$  denoising. We solve (3.1) using the split Bregman algorithm, see [9, Chapter 4] for more details.

First we present some one dimensional examples that verify numerically our analytical results. Figure 2a shows the  $\text{TVL}^\infty$  result for the function  $g(x) = h\mathcal{X}_{(0,L)}(x) + \lambda x$  where  $\alpha, \beta$  belong to the yellow region of Figure 1b. Note that the numerical and the analytical results coincide. We have also computed the TGV solution where the parameters are selected so that  $\|f - u_{\text{TGV}}\|_2 = \|f - u_{\text{TVL}^\infty}\|_2$ . Figure 2b shows a numerical verification of Proposition 4. There, the  $\text{TVL}^\infty$  parameters are  $\alpha = 2$  and  $\beta = 4$ , while the TGV ones are  $\alpha = 2$  and  $\beta = 2$ . Both solutions coincide since they satisfy the symmetry properties of the proposition.

We proceed now to two dimensional examples, starting from Figure 3. There we used a synthetic image corrupted by additive Gaussian noise of  $\sigma = 0.01$ , cf.

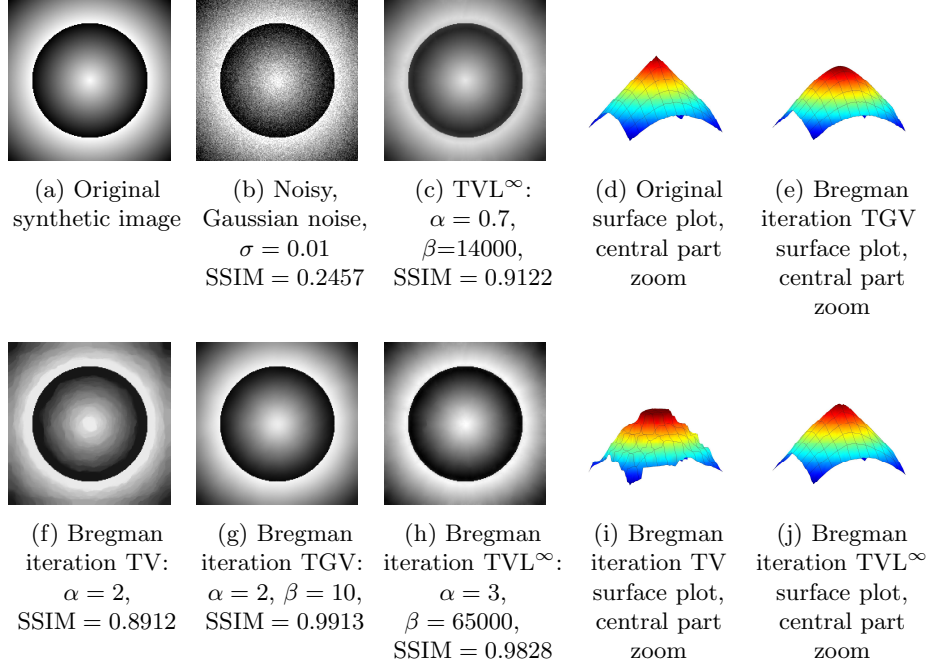


Fig. 3:  $\text{TVL}^\infty$  based denoising and comparison with the corresponding TV and TGV results. All the parameters have been optimised for best SSIM.

Figures 3a–3b. We observe that  $\text{TVL}^\infty$  denoises the image in a staircasing-free way in Figure 3c. In order to do so however, one has to use large values of  $\alpha$  and  $\beta$  something that leads to a loss of contrast. This can easily be treated by solving the *Bregman iteration* version of  $L^2$ - $\text{TVL}_{\alpha,\beta}^\infty$  minimisation, that is

$$\begin{aligned} u^{k+1} &= \operatorname{argmin}_{u,w} \frac{1}{2} \|f - v^k - u\|_2^2 + \alpha \|\nabla u - w\|_1 + \beta \|w\|_\infty, \\ v^{k+1} &= v^k + f - u^{k+1}. \end{aligned} \quad (4.1)$$

Bregman iteration has been widely used to deal with the loss of contrast in these type of regularisation methods, see [2, 7] among others. For fair comparison we also employ the Bregman iteration version of TV and TGV regularisations. The Bregman iteration version of  $\text{TVL}^\infty$  regularisation produces visually a very similar result to the Bregman iteration version of TGV, even though it has a slightly smaller SSIM value, cf. Figures 3g–3h. However,  $\text{TVL}^\infty$  is able to reconstruct better the sharp spike at the middle of the figure, cf. Figures 3d, 3e and 3j.

The reason for being able to obtain good reconstruction results with this particular example is due to the fact that the modulus of the gradient is essentially constant apart from the jump points. This is favoured by the  $\text{TVL}^\infty$  regularisation which promotes constant gradients as it is proved rigorously in dimension one in Proposition 3. We expect that a similar analytic result holds in higher

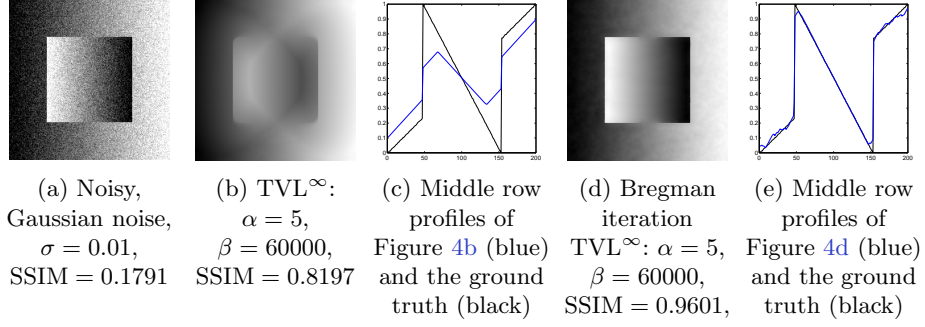


Fig. 4: Illustration of the fact that  $\text{TVL}^\infty$  regularisation favours gradients of fixed modulus

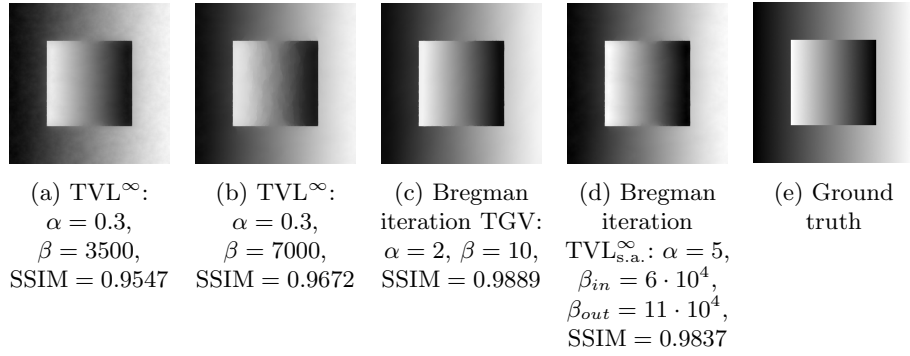


Fig. 5:  $\text{TVL}^\infty$  reconstructions for different values of  $\beta$ . The best result is obtained by spatially varying  $\beta$ , setting it inversely proportional to the gradient, where we obtain a similar result to the TGV one

dimensions and we leave that for future work. However, this is restrictive when gradients of different magnitude exist, see Figure 4. There we see that in order to get a staircasing-free result with  $\text{TVL}^\infty$  we also get a loss of geometric information, Figure 4b, as the model tries to fit an image with constant gradient, see the middle row profiles in Figure 4c. While improved results can be achieved with the Bregman iteration version, Figure 4d, the result is not yet fully satisfactory as an *affine staircasing* is now present in the image, Figure 4e.

**Spatially adapted  $\text{TVL}^\infty$ :** One way to allow for different gradient values in the reconstruction, or in other words allow the variable  $w$  to take different values, is to treat  $\beta$  as a spatially varying parameter, i.e.,  $\beta = \beta(x)$ . This leads to the spatially adapted version of  $\text{TVL}^\infty$ :

$$\text{TVL}_{\text{s.a.}}^\infty(u) = \min_{w \in L^\infty(\Omega)} \alpha \|Du - w\|_{\mathcal{M}} + \|\beta w\|_{L^\infty(\Omega)}. \quad (4.2)$$

The idea is to choose  $\beta$  large in areas where the gradient is expected to be small and vice versa, see Figure 5 for a simple illustration. In this example the



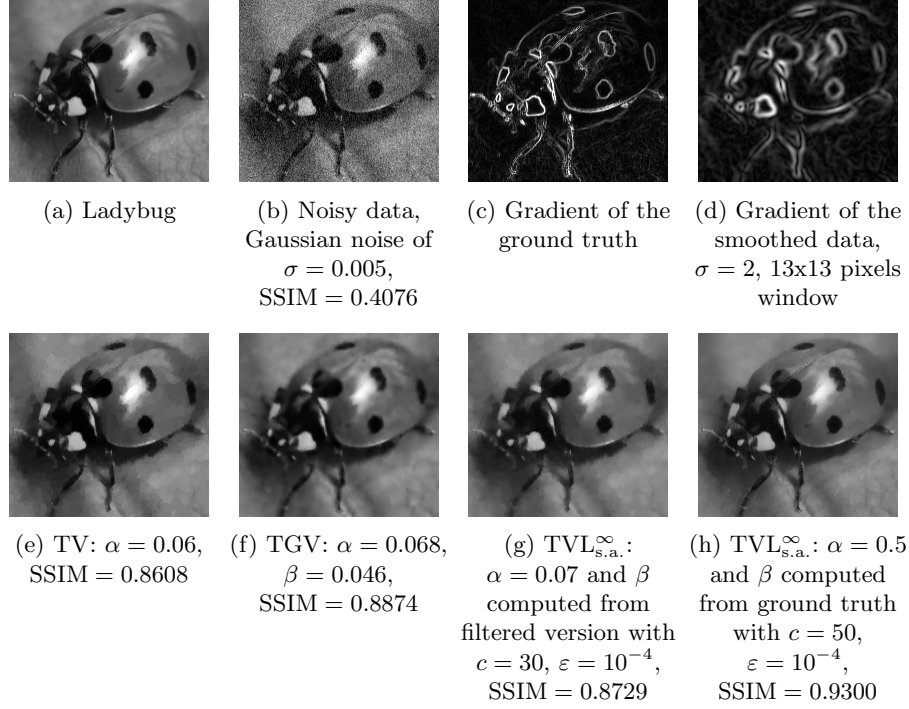


Fig. 6: Best reconstruction of the “Ladybug” in terms of SSIM using TV, TVG and spatially adapted  $\text{TVL}^\infty$  regularisation. The  $\beta$  for the latter is computed both from the filtered (Figure 6g) and the ground truth image (Figure 6h)

slope inside the inner square is roughly twice the slope outside. We can achieve a perfect reconstruction inside the square by setting  $\beta = 3500$ , with artefacts outside, see Figure 5a and a perfect reconstruction outside by setting the value of  $\beta$  twice as large, i.e.,  $\beta = 7000$ , Figure 5b. In that case, artefacts appear inside the square. By setting a spatially varying  $\beta$  with a ratio  $\beta_{\text{out}}/\beta_{\text{in}} \simeq 2$  and using the Bregman iteration version, we achieve an almost perfect result, visually very similar to the TGV one, Figures 5c–5d. This example suggests that ideally  $\beta$  should be inversely proportional to the gradient of the ground truth. Since this information is not available in practice we use a pre-filtered version of the noisy image and we set

$$\beta(x) = \frac{c}{|\nabla f_\sigma(x)| + \epsilon}. \quad (4.3)$$

Here  $c$  is a positive constant to be tuned,  $\epsilon > 0$  is a small parameter and  $f_\sigma$  denotes a smoothing of the data  $f$  with a Gaussian kernel. We have applied the spatially adapted  $\text{TVL}^\infty$  (non-Bregman) with the rule (4.3) in the natural image “Ladybug” in Figure 6. There we pre-smooth the noisy data with a discrete Gaussian kernel of  $\sigma = 2$ , Figure 6d, and then apply  $\text{TVL}_{\text{s.a.}}^\infty$  with the rule (4.3), Figure 6g. Comparing to the best TGV result in Figure 6f, the SSIM value is slightly smaller but there is a better recovery of the image details (objective). Let

us note here that we do not claim that our rule for choosing  $\beta$  is the optimal one. For demonstration purposes, we show a reconstruction where we have computed  $\beta$  using the gradient of ground truth  $u_{\text{g.t.}}$ , Figure 6c, as  $\beta(x) = c/(|\nabla u_{\text{g.t.}}(x)| + \epsilon)$ , with excellent results, Figure 6h. This is of course impractical, since the gradient of the ground truth is typically not available but it shows that there is plenty of room for improvement regarding the choice of  $\beta$ . One could also think of reconstruction tasks where a good quality version of the gradient of the image is available, along with a noisy version of the image itself. Since the purpose of the present paper is to demonstrate the capabilities of the  $\text{TVL}^\infty$  regulariser, we leave that for future work.

**Acknowledgments.** The authors acknowledge support of the Royal Society International Exchange Award Nr. IE110314. This work is further supported by the King Abdullah University for Science and Technology (KAUST) Award No. KUK-I1-007-43, the EPSRC first grant Nr. EP/J009539/1 and the EPSRC grant Nr. EP/M00483X/1. MB acknowledges further support by ERC via Grant EU FP 7-ERC Consolidator Grant 615216 LifeInverse. KP acknowledges the financial support of EPSRC and the Alexander von Humboldt Foundation while in UK and Germany respectively. EP acknowledges support by Jesus College, Cambridge and Embiricos Trust.

## References

1. L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems*. Oxford University Press, USA, 2000.
2. M. Benning, C. Brune, M. Burger, and J. Müller. Higher-order TV methods – Enhancement via Bregman iteration. *J. Sci. Comput.*, 54(2-3):269–310, 2013.
3. K. Bredies, K. Kunisch, and T. Pock. Total generalized variation. *SIAM J. Imaging Sci.*, 3(3):492–526, 2010.
4. K. Bredies, K. Kunisch, and T. Valkonen. Properties of  $L^1$ -TGV<sup>2</sup> : The one-dimensional case. *Journal of Mathematical Analysis and Applications*, 398(1):438 – 454, 2013.
5. M. Burger, K. Papafitsoros, E. Papoutsellis, and C.-B. Schönlieb. Infimal convolution regularisation functionals of BV and  $L^p$  spaces. Part I: The finite  $p$  case. *submitted, arXiv:1504.01956*.
6. P. Huber. Robust regression: Asymptotics, conjectures and monte carlo. *The Annals of Statistics*, 1:799–821, 1973.
7. S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Model. Simul.*, 4:460–489.
8. K. Papafitsoros and K. Bredies. A study of the one dimensional total generalised variation regularisation problem. *Inverse Probl. Imaging*, 9(2):511–550, 2015.
9. E. Papoutsellis. *First-order gradient regularisation methods for image restoration. Reconstruction of tomographic images with thin structures and denoising piecewise affine images*. PhD thesis, University of Cambridge, 2015.
10. L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Phys. D*, 60(1-4):259–268.